

PCT09

RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/913,767

TIME: 14:17:50

Input Set : A:\Gks1001.app

Output Set: N:\CRF3\09102001\I913767.raw

3 <110> APPLICANT: Frommer, Wolf G.

5 <120> TITLE OF INVENTION: Nucleic Acids That Code for a Nucleobase Transporter

7 <130> FILE REFERENCE: GKS-101.0 7911/83549

C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/913,767

10 <141> CURRENT FILING DATE: 2001-08-17

12 <150> PRIOR APPLICATION NUMBER: PCT/EP00/01397

13 <151> PRIOR FILING DATE: 2000-02-21

15 <150> PRIOR APPLICATION NUMBER: DE 19907209.4

16 <151> PRIOR FILING DATE: 1999-02-19

18 <160> NUMBER OF SEQ ID NOS: 13

20 <170> SOFTWARE: PatentIn Ver. 2.1

22 <210> SEQ ID NO: 1

23 <211> LENGTH: 1225

24 <212> TYPE: DNA

25 <213> ORGANISM: Arabidopsis thaliana

27 <400> SEQUENCE: 1

28 aaaacagcaa gcagcaagaa gaagatgaag aatggtttga taatcataaa ctgtattatc 60
 29 ctactatag gaacatgtgg aggtcctttg ttaactcgtc tctacttcac caatggcgga 120
 30 aaacgaatct ggttcatgag ctctctatca accgctggtt ttccaatcat cctcatccct 180
 31 ctcttggtct ccttcctcag cgtcgcgcgc ggcaaccgca accctaacia cgcggaaaac 240
 32 aagcggaaaa caaagctctt cctcatggaa actcctctgt ggagcgcttc cattgtcata 300
 33 gggttgctca caggacttga caactactta tattcttaag gattagcata tctgccagtt 360
 34 tcaacttcat cgctcataat cggaactcaa ctactttca acgctctctt cgctttcttg 420
 35 ttagtcaagc aaaagttcac tccgttctcc ataaacgccg tctgtttgtt gacggtttgt 480
 36 atcgggatcc ttgcgttaca cagtgatgga gacaaaccgg ctaaggagag caagaaagag 540
 37 tatgtggttg ggttcttgat gactgtggtt gcagctcttc tctatgcttt tatattaccg 600
 38 ctggttgagc taacttacaa gaaagctcgt caagaaatca ctttccact tgtgcttgag 660
 39 attcagatgg tcatgtgctt tgctgctact ttttctgtg tcattggcat gttcatcggt 720
 40 ggagatttta aggtgatagc aagagaagca agagagttca agattggagg atcagtggtt 780
 41 tactatgcat tgatagtgat cacaggaata atatggcaag gtttcttctt aggagccata 840
 42 gggatttgtt tttgtgcatc atcactagct tctggtgttc tgataagtgt tctgcttcgg 900
 43 gtgactgaag ttttcgccgt cgtttgtttc cgggagaagt ttcaggcaga gaaagggtgc 960
 44 tctctacttc tttctctttg gggatttgtc tcttacttct acggcgagtt taaatccggc 1020
 45 aagaaagtgt ttgataaacc tcaaccgccg gagacagaac tgcctattct tccagttagt 1080
 46 gattatgttg cttaatttct ataactctat acgattataa cagagcatta ctgttatgtt 1140
 47 ttgttcctaa atattatgtg tgattgtgtg tttttgttat tgttcttggt tataagtatg 1200
 48 aataaaattt gaaagatatg gagct 1225

51 <210> SEQ ID NO: 2

52 <211> LENGTH: 1049

53 <212> TYPE: DNA

54 <213> ORGANISM: Arabidopsis thaliana

56 <400> SEQUENCE: 2

57 aagatgaaga tgaagacagt tcttgtaatc ataaactgta tattcttggc cattggaaac 60
 58 tgtggaggcc ctctaattgat gcgtctctac ttccaaaatg gtggcgaaag gatctggttt 120
 59 ccaagcttcc ttcaaaccgt tgggtgtcca ctcatcttct tccctcttct cttatctttc 180
 60 ctccgccgct gtcgttgctt tgaagaacaa gaaacgactc catttttctt catgaaacct 240
 61 cctctcttta tcgccgctat cgttggttgtt ttgctcgtgg gatttgacaa ttacctctac 300

ENTERED

p. 5

RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/913,767

TIME: 14:17:50

Input Set : A:\Gks1001.app

Output Set: N:\CRF3\09102001\I913767.raw

```

62 tcttacgggt tagcttatat ccctgtttct actgctctt tgatcatctc cgcgcaatta 360
63 ggcttcactg ctctctttgc attttttatg gtgaagcaaa agttcacacc tttcactata 420
64 aacgctatcg ttttgctcac tgggtggtgcc gtagtccttg cccttaactc tgatagtgc 480
65 aagcttgcaa acgagacaca caaggaatat gttgttgggt tcatcatgac tcttgggtgca 540
66 gctcttctct atgggtttat attgccactt gtcgagcttt cttacaagaa atctgggtcag 600
67 cgaatcacgt atacgctcgc gctcgagtgc cagatggtct tatgctttgc tgccacttgt 660
68 gtctgcctcg tggggatgct agccgctggc gatttcaagg tgatagcagg agaagcaaga 720
69 gattttaagc ttggagagtc tttgtactat gtggtgattg tgttcacggc cataatctgg 780
70 caagcatttt ttgtgggagc tattggggtg atcttctgtg catcgtctct ggtctctgga 840
71 attatggtca gtgctctgct tccggtgacg gtgatcttgg ccgtcatttg cttccaggag 900
72 aagtttcagg cggggaaagg tgctgctttg gctctctccc tctggggatc agtctcttat 960
73 ttctatggac aggttaaate cgaggagaag actaaggctc aggatacaca actgtctcag 1020
74 cttccagtta ctgattatgt agcttaaaa 1049
77 <210> SEQ ID NO: 3
78 <211> LENGTH: 1145
79 <212> TYPE: DNA
80 <213> ORGANISM: Arabidopsis thaliana
82 <400> SEQUENCE: 3
83 ctgtttcatg tgttgatggt agaacctgaa gggaaatfff caacagaaga gagaagtcac 60
84 aagtactctt ggaggttaag agtgctcttc tatgtcactc tctcttagc tggagagaca 120
85 atagccactc tcttaggtag actttactac gaaaaaggcg gtaaaagcac atggctcgaa 180
86 accttggttc agctttagg gtttctctta acccttctct gctattatta cttaaagcct 240
87 gagecgtcca agactaaaac cattaccaa aaaactactt ctctctctt gacactatct 300
88 ttagtgtata ttgacttgg cttgcttgtt gctggacatt gtattttgta ctcatttggg 360
89 ctactttacc ttcctgtctc aactttctct ttgatctctg cgtcgcaatt ggcttttaac 420
90 gccgtcttct cttacttctt aaactcaca aaaatcacac catttatact caattcactt 480
91 gttctcttaa ccatatcttc tacacttctt gttatccaac atgaaccaga atctccctct 540
92 tctacttcaa agtccgcagc caagtccaag tatgtgattg gatacatctg cgcggtcggg 600
93 agctcagctg gttattctct ggtgctttct ttaacagatt acgcgttcga aaagattcta 660
94 aagaaataca cattcaaggc tattttagac atggccacat atccgtctat ggtagctact 720
95 tgtgtagttg tggtaggact ttttggaagt ggtgggtgga aaaagctgag tacagaaatg 780
96 gaagagtttc aactagggaa aagctcatac attttgataa acatcggttc aacgatatac 840
97 tggcaagctt gtttgattgg aagtgttgg ttgattatcg aagtttcac gcttttttcc 900
98 aatgtcataa gcactctttg tttaccagtt gtgectgttc ttgctgttgt cttcttccgt 960
99 gatgagatga gtggaatcaa gttggttgca atgtttttg ccatctgggg atttgtttct 1020
100 tatggttatc agcattatgt caatgataga aagccagaag aagaccaaga gcttctcag 1080
101 tctaaagaag aagaagaaca aaaacaagta gataccattc atgtccaagc ttaggcaaaag 1140
102 atcca 1145
105 <210> SEQ ID NO: 4
106 <211> LENGTH: 1293
107 <212> TYPE: DNA
108 <213> ORGANISM: Arabidopsis thaliana
110 <400> SEQUENCE: 4
111 ggaagtctc ttagagtggg catggaaata actcaagtaa tctatgtcaa tggtaagcaa 60
112 gatgcatctc gtagttaga ttacttgatt cttttcgcta acctgttgtt ttgatgttt 120
113 tcagatcata acaacataga agcaaaccta acaggtcagg aggaaatgaa taccaccatg 180
114 gaaatcgaat ctctgtccgt acctcaatcg aagaactata agaaatggct tcgtatttcc 240
115 atttacgtgt tctttgtcct tgcttgccaa gcactttcta caattttggg cagagtttac 300
116 tatgaaaatg gtgggaagag tacatggatg ggaacacttg tccaactaat cggcttccct 360

```

RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/913,767

TIME: 14:17:50

Input Set : A:\Gks1001.app

Output Set: N:\CRF3\09102001\I913767.raw

```

117 gttctgtttc tcttccgctt cttttcccaa accaaaaatc ccaaaccaac agaagcagat 420
118 ttcagaaagt tctcttcctt caccattctt ggatcagttt acatcgttac tggactatta 480
119 gtgtctgcta actcttatat gtctctgtt ggtttactat acttaccagt ttctactttc 540
120 tccctcatct tggcctcaca attggccttc actgccttct tctcatattt tctaaactcg 600
121 cagaagttca cacctttcat tgtgaattct ctgtttctcc ttactatttc ctctgccctc 660
122 ctctgtgttca aactgattc ggaaaacaca gcaaaagtgt ctagagtaaa atatgtgata 720
123 gggataatat gtaccattgg tgcttctgct gggattggat tgctgctatc cctggtacaa 780
124 ctgactctca ggaaggtttt aaagaagcaa acattctcaa cggtcactga cttggtcgct 840
125 taccaatctc tagttgcaag ctgtgtggtt ctcataggac ttttcgcaag cggggagtgg 900
126 aaaactttta caagtggagt ggaaaactac aaactgggga aagtgccata cgttatgact 960
127 ttggcctcga tagctatttc ctggcaagtc tacaccattg gcgtcgtggg actgatcttt 1020
128 gagtcatctt ctgtgttctc caattccata actgctgtgg gattgcctat agttccagtt 1080
129 gtagcagtga ttgttttcca tgataaaatg aacgcgtcaa agatcttctc catcatttta 1140
130 gctatctggg gattcatttc atttgtctat cagcactacc tcgacgaaaa gaagttgaag 1200
131 actagccaca caagtctgt aggagatcct catctactac ctgctgagga aggtcacaca 1260
132 aacatacata gtgtatgata aaaacatatt tcc 1293
135 <210> SEQ ID NO: 5
136 <211> LENGTH: 1194
137 <212> TYPE: DNA
138 <213> ORGANISM: Arabidopsis thaliana
140 <400> SEQUENCE: 5
141 tcatgagata taataaacat gagtgttaat ttttcagggtg accagaactt agaagcaaac 60
142 cttatagatc atgagggtgt aactgaatca tcatcatcag ctgtgcctca aaccgagaac 120
143 tataaaagggt ggcttcgtgt ctccatatac gtaatctttg tcctcttttg ccagccacta 180
144 gctacaattc tgggtagatt gtactatgaa aatggaggaa atagcacata tgtggtaaac 240
145 cttcttcaac tcatggctt cctgtactg gttctgttcc gcttcttttc tcgaatcagg 300
146 caacccaaat caacagatac aaatttccagt cagtccctt ccttcaccac ctttgcctcg 360
147 gtttacttgt gcaactggact gctagtgtcc gcttatgctt atttgtctgc agtaggggtg 420
148 ctctacttac cagtctctac tttctccctc atcttggcct cacagttggc cttcactgcc 480
149 tttttctcat atttctttaa ctgcgaaaag ttcaactcct tgatagtcag ttctttgctt 540
150 ctctcactg tatcctctgc tcttcttggt gtcaaacactg attcagaaaa ctcaactaat 600
151 gtatctagag tacagtatgt gatcggttcc atatgtacca tcggtgcttc cgctgggatt 660
152 ggactgttac tatctctgat acaaatgctc ttcaggaaaag ttttcacgaa gcatacatoc 720
153 tcagcagtca cggacttggc catttaccag tctctagttg cgagttgtgt agttctcata 780
154 ggactttttg caagtggaga gtgggaaact ttgccaagtg agatgagaaa ctacaaactc 840
155 gggaaagtgt catatgtttt gacttttagcc tcggcagcta tttcctggca agtctacact 900
156 cctggtcttg tgggattgat cttcgagtca tctctgtgt tctccaattc cataacagct 960
157 gtgggattgc ctatagttcc agttgcggca gtgatagttt tccatgatag aatggacgca 1020
158 tccaaaatct tctccattat tttagctatc tgcggttcc tttcattcgt ctatcagcac 1080
159 tacctcgacg aaaagaagtt gaatactagc cacacaagtg ctgtaggaga tcttcatcta 1140
160 cctgttgagg aaggtcacac aaacatacaa agtgtgtgat caaagcatat ttcc 1194
163 <210> SEQ ID NO: 6
164 <211> LENGTH: 1081
165 <212> TYPE: DNA
166 <213> ORGANISM: Arabidopsis thaliana
168 <400> SEQUENCE: 6
169 caaatccaac agttcaagat gaaagaaatt cagtcagtag cagccaagca gaagtatctc 60
170 actetaacac atacaaacgg tggctcagga gtatactatg acaacggagg aaacagtaaa 120
171 tggttagcaa cggtagttca acttgttggc tttcctgtgc tacttccata ttatatcttg 180

```

RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/913,767

TIME: 14:17:50

Input Set : A:\Gks1001.app

Output Set: N:\CRF3\09102001\I913767.raw

```

172 tcattttaaaa cacatgcaac aactgataga gatggaaaaa gaacctcacc taggaaccgt 240
173 gtattggttt acgtagtgtt tggacttctt gtaggagcag attgctatct gtactccatt 300
174 ggacttcttt acttaccctt ttctacctat tccctgatct gtgcatctca gttagecctt 360
175 aatgctttct tctcttattt tcttaactca caaaaactta cccctatcat tttaaattct 420
176 cttttcctct ttaactatct ttccacccta cttgcattca ataatgagga gacagactcc 480
177 acaaaagtta caaaaggaga gtatgtcaaa ggtttcatat gcaccgttgc tgcgtctgct 540
178 ggttatggtc tagtcttctt cctacaacag ctagcctttc taaaagtcct aaagaagcaa 600
179 aatttctcag aagttatgga tatgataatc tacgtgagtc tagtggccag ttgtgttagc 660
180 gtgggtggggc tttttgctag cagtgaagtgg aaaactttga gcagtgaat ggataactac 720
181 aaacatggga aggtatccta cattatgaac ctagtgtgga cagctgttac ctggcagtta 780
182 ttctccatcg gtggcacagg actgatcttc gagctctctt ctctattctc aaatgcaata 840
183 agcgttttgg gactcccagt ggttcctatc ttggctgtaa tcattttcca tgacaaaatg 900
184 aatgggttaa aggtgatttc tatgatccta gctatttggg gtttcacttc ctatgtctac 960
185 caacaatctt ttgatgacaa aaacttgaag aaaaatcatg aaatcacac aacagaatcc 1020
186 cctgaccac cagaagcaga agagtcaact tggcaatcaa aataagctga tattttgaaa 1080
187 g 1081
190 <210> SEQ ID NO: 7
191 <211> LENGTH: 1071
192 <212> TYPE: DNA
193 <213> ORGANISM: Arabidopsis thaliana
195 <400> SEQUENCE: 7
196 gagggggatt ccacatctac tatgaagggg gatcaagaag tacaagtcac tggccaatca 60
197 gttgctacaa ttctgggcag actatactat gaaaatggag gaaacagcaa atggctagca 120
198 acggtagttc agcttgtagg ctttcctatt ctacttccat atcatctctt gtcagtcaaa 180
199 acacatacaa caactcagag agatggcaaa ttaacctcac ttaggaaccg tgcattagtt 240
200 tacatagtgc ttggacttct tgtaggagca gcttgctacc tatattccat tggactgctt 300
201 tacctacctg tttctacctt ttccctgata tgtgcatcac agttagcctt caccgctttc 360
202 ttctcttatt tactcaactc acaaaaactt actcctatca ttttgaattc tcttttcttc 420
203 ctactatat ctccaccct ccttgcattt aataacgagg aatcagattc caaaaaagt 480
204 acaaaaggag agtatgtcaa aggtttcgta tgcaccgttg gtgcatctgc tgggtttgg 540
205 ctactcttat ccctacaaca gctagccttt cgtaaagttt taaagaagca aactttctca 600
206 gaagtataaa atatgataat ctacatgagt ctagtggcca gttgtgttag cgtgggtggg 660
207 ctttttgcta gtagcgagtg gaaaactttg agcagtgaat tggaaaacta caaacttggg 720
208 aaggtatcct atgtcatgaa cctagtgttg acagctgtta cctggcaggg attctccatc 780
209 ggttgcacag gactgatctt cgagctttcc tccctattct caaatgcaat aagcgcttg 840
210 ggactccccg tggttcctat cctggctgtc atcattttcc atgacaaaat gaacggctta 900
211 aaggtgattt ctatgattct agctatttgg ggtttcgta cctatgtcta ccaacaatat 960
212 cttgatgaaa caaacttgaa gaaaagtaat gaaataccaa caacagaatc ccctgaccga 1020
213 ccagaagcag aaggggtcaag tgagcaatca aaataagctg ttacttcaaa g 1071
216 <210> SEQ ID NO: 8
217 <211> LENGTH: 356
218 <212> TYPE: PRT
219 <213> ORGANISM: Arabidopsis thaliana
221 <400> SEQUENCE: 8
222 Met Lys Asn Gly Leu Ile Ile Asn Cys Ile Ile Leu Thr Ile Gly
223 1 5 10 15
225 Thr Cys Gly Gly Pro Leu Leu Thr Arg Leu Tyr Phe Thr Asn Gly Gly
226 20 25 30
228 Lys Arg Ile Trp Phe Met Ser Phe Leu Ser Thr Ala Gly Phe Pro Ile

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/913,767

DATE: 09/10/2001
 TIME: 14:17:50

Input Set : A:\Gks1001.app
 Output Set: N:\CRF3\09102001\I913767.raw

```

229          35          40          45
231 Ile Leu Ile Pro Leu Leu Val Ser Phe Leu Ser Arg Arg Arg Gly Asn
232          50          55          60
234 Arg Asn Pro Asn Asn Ala Glu Asn Lys Arg Lys Thr Lys Leu Phe Leu
235 65          70          75          80
237 Met Glu Thr Pro Leu Phe Ile Ala Ser Ile Val Ile Gly Leu Leu Thr
238          85          90          95
240 Gly Leu Asp Asn Tyr Leu Tyr Ser Tyr Gly Leu Ala Tyr Leu Pro Val
241          100          105          110
243 Ser Thr Ser Ser Leu Ile Ile Gly Thr Gln Leu Ala Phe Asn Ala Leu
244          115          120          125
246 Phe Ala Phe Leu Leu Val Lys Gln Lys Phe Thr Pro Phe Ser Ile Asn
247          130          135          140
249 Ala Val Val Leu Leu Thr Val Gly Ile Gly Ile Leu Ala Leu His Ser
250 145          150          155          160
252 Asp Gly Asp Lys Pro Ala Lys Glu Ser Lys Lys Glu Tyr Val Val Gly
253          165          170          175
255 Phe Leu Met Thr Val Val Ala Ala Leu Leu Tyr Ala Phe Ile Leu Pro
256          180          185          190
258 Leu Val Glu Leu Thr Tyr Lys Lys Ala Arg Gln Glu Ile Thr Phe Pro
259          195          200          205
261 Leu Val Leu Glu Ile Gln Met Val Met Cys Leu Ala Ala Thr Phe Phe
262          210          215          220
264 Cys Val Ile Gly Met Phe Ile Val Gly Asp Phe Lys Val Ile Ala Arg
265 225          230          235          240
267 Glu Ala Arg Glu Phe Lys Ile Gly Gly Ser Val Phe Tyr Tyr Ala Leu
268          245          250          255
270 Ile Val Ile Thr Gly Ile Ile Trp Gln Gly Phe Phe Leu Gly Ala Ile
271          260          265          270
273 Gly Ile Val Phe Cys Ala Ser Ser Leu Ala Ser Gly Val Leu Ile Ser
274          275          280          285
276 Val Leu Leu Pro Val Thr Glu Val Phe Ala Val Val Cys Phe Arg Glu
277          290          295          300
279 Lys Phe Gln Ala Glu Lys Gly Val Ser Leu Leu Ser Leu Trp Gly
280 305          310          315          320
282 Phe Val Ser Tyr Phe Tyr Gly Glu Phe Lys Ser Gly Lys Lys Val Val
283          325          330          335
285 Asp Lys Pro Gln Pro Pro Glu Thr Glu Leu Pro Ile Leu Pro Val Ser
286          340          345          350
288 Asp Tyr Val Ala
289          355
292 <210> SEQ ID NO: 9
293 <211> LENGTH: 352
294 <212> TYPE: PRT
295 <213> ORGANISM: Arabidopsis thaliana
297 <400> SEQUENCE: 9
298 Met Val Lys Ala Leu Val Ile Ile Asn Cys Ile Ile Leu Ala Ile Gly
299 1          5          10          15
301 Asn Cys Gly Gly Pro Leu Ile Met Arg Leu Tyr Phe Asn Asn Gly Gly

```

Use of n and / or Xaa has been detected in the
 Sequence Listing. Review the Sequence Listing
 to ensure a corresponding explanation is present
 in the <220> to <223> fields of each sequence
 using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/913,767

DATE: 09/10/2001

TIME: 14:17:51

Input Set : A:\Gks1001.app

Output Set: N:\CRF3\09102001\I913767.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number
L:433 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:474 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:477 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:510 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13